1 Bait construct

DNA binding domain (BD)

ANK fragment ("bait")

Plasmid DNAs

2 Mouse kidney cDNA library

"prey"   Activation domain (AD)

P

Q

R

S
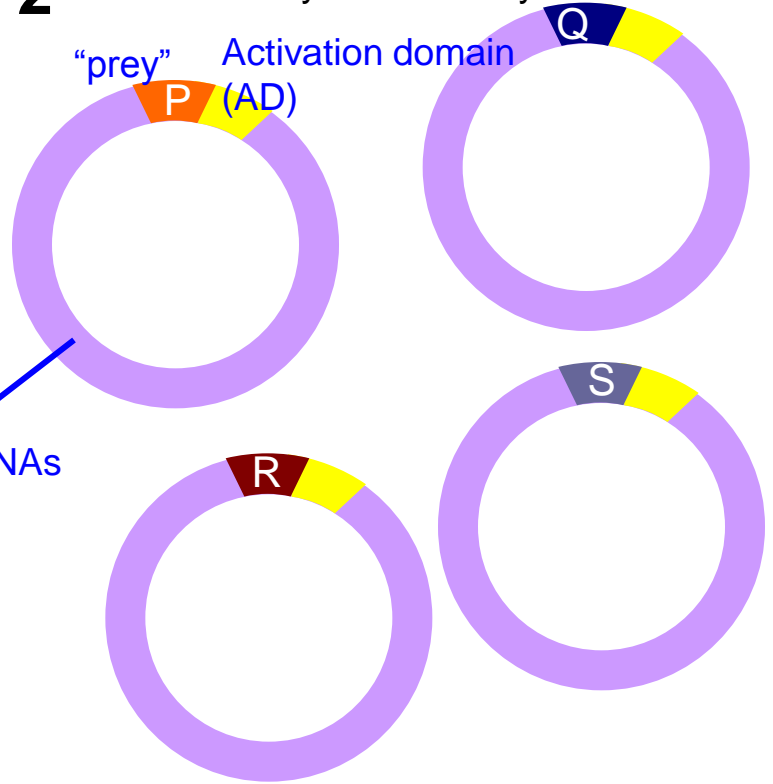
Transform bait and prey plasmid combinations into yeast cells

3 Interaction of the bait and prey proteins

BD   ANK B   P   AD
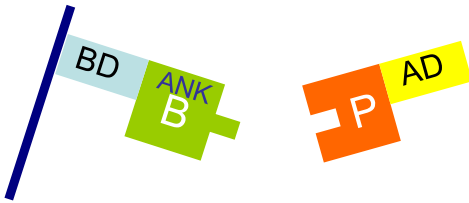
4 BD is now indirectly connected to AD, allowing activation of transcription of the reporter gene

BD   ANK B   P   AD

5

No His3 product

His3 product

Reporter gene product is only formed if bait and prey proteins interact

Grow yeast cells in histidine-deficient medium

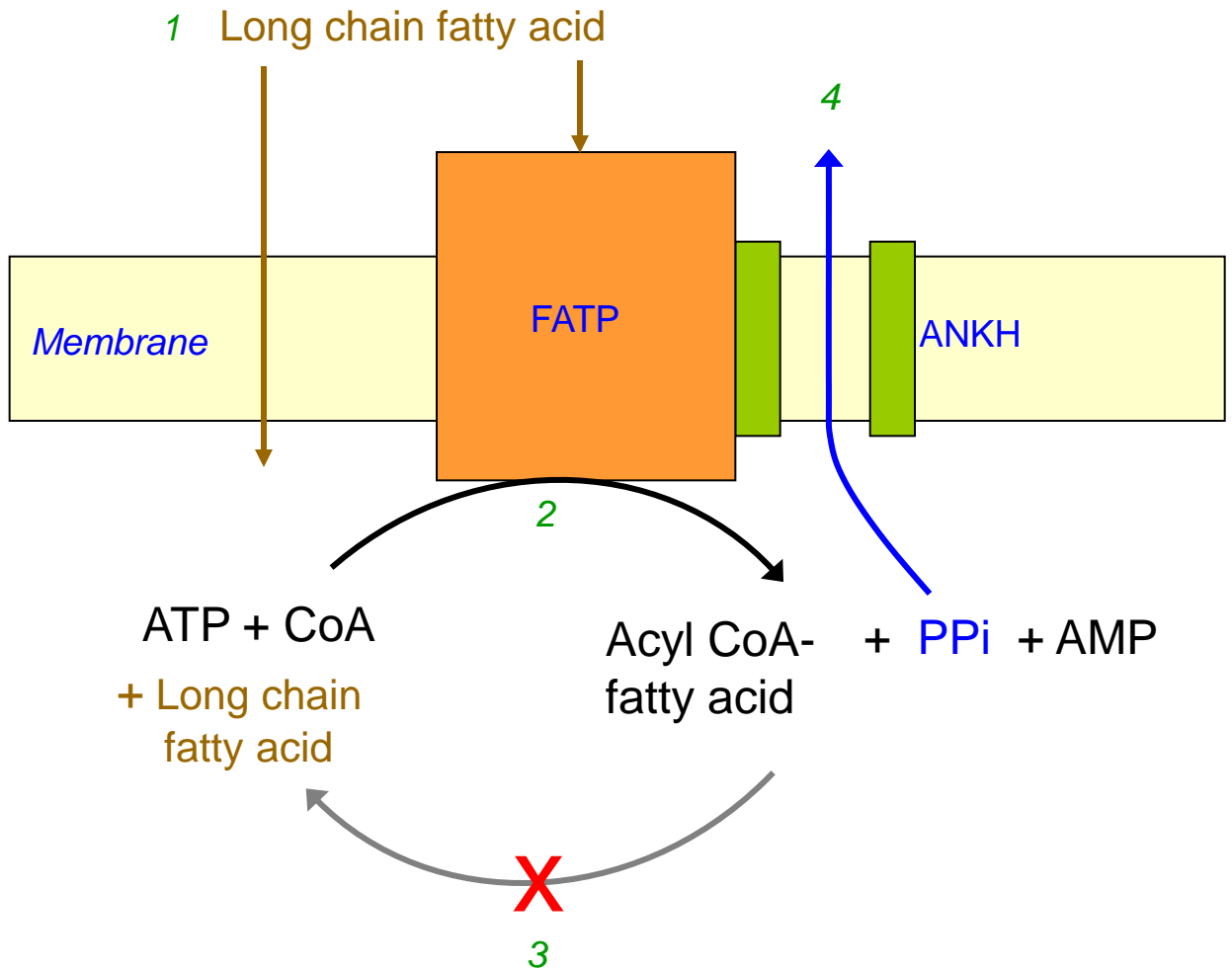No growth

Purify and sequence DNA

72

Figure 5.14 Possible interaction between FATP and ANKH. Long chain fatty acids diffuse or are transported across the membrane by FATP (**1**). An associated acyl CoA synthetase catalyses the esterification of the fatty acid (**2**). Energy comes from ATP, and produces PPi. To prevent reversal of this reaction (**3**), PPi must be immediately removed, possibly by compartmentalisation across the membrane via ANKH (**4**).
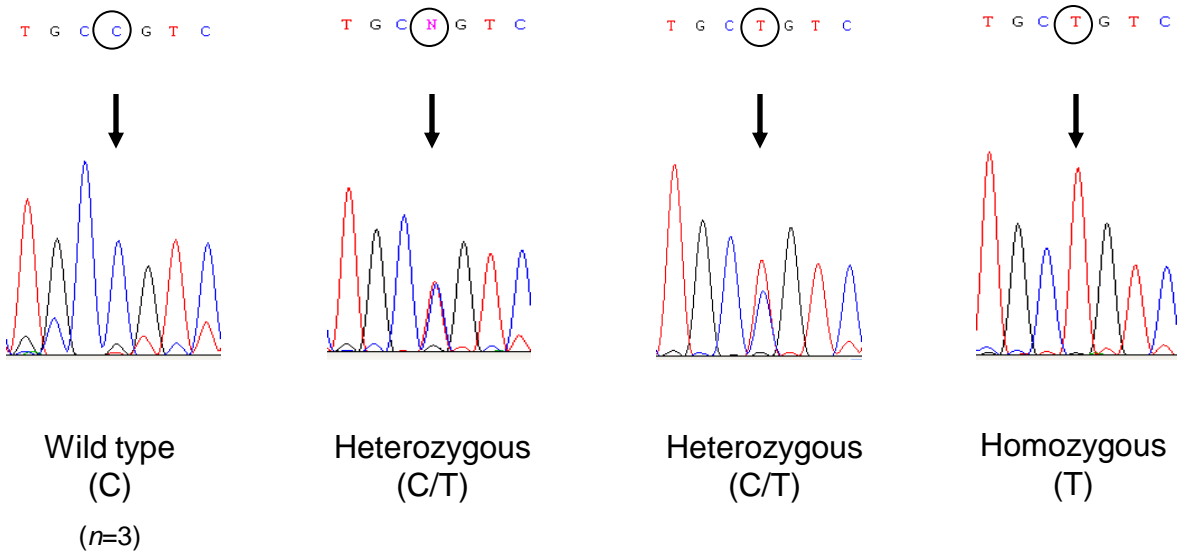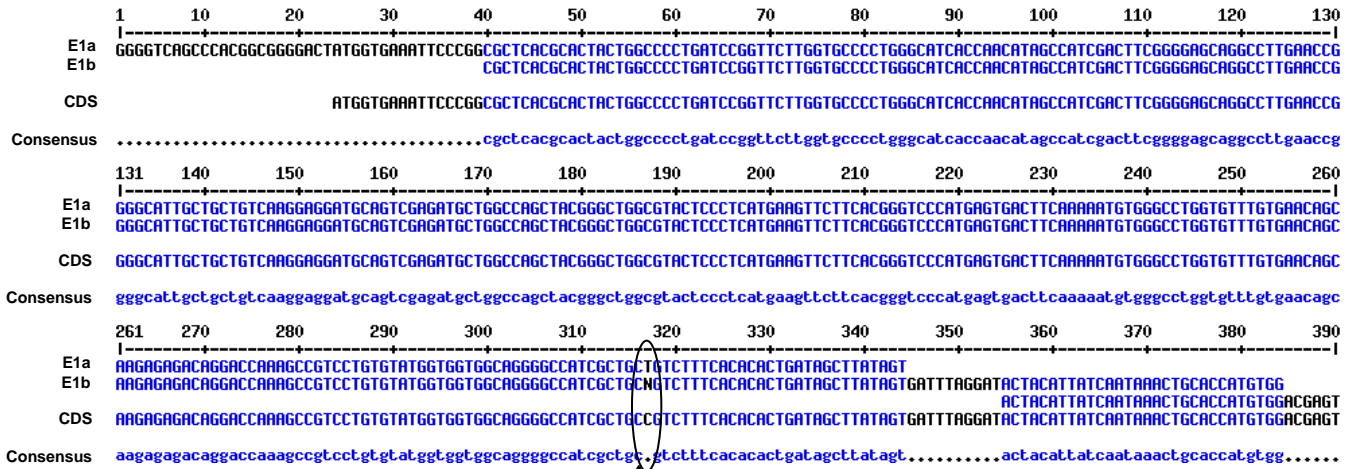
Figure 3.7 Detection of a single nucleotide polymorphism in ANKH coding sequence in six human subjects. Part of the contiguous nucleotide sequence for one of the subjects is shown (above). This fragment was sequenced twice (E1a, E1b) using different primer sets, and is aligned with the ANKH coding sequence NM054027. Extracts from sequencing chromatograms (below) show that the nucleotide at position 294 (317 on sequence above) was either a cytosine (C) or thymine (T), or both alleles were present. The resulting codons GCC or GCT are synonymous for alanine at residue position 97. This SNP is annotated as RS17251667.
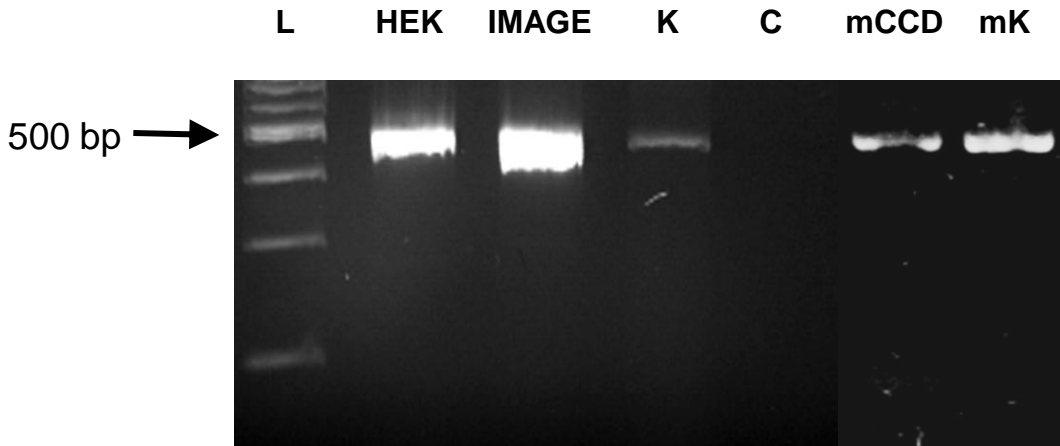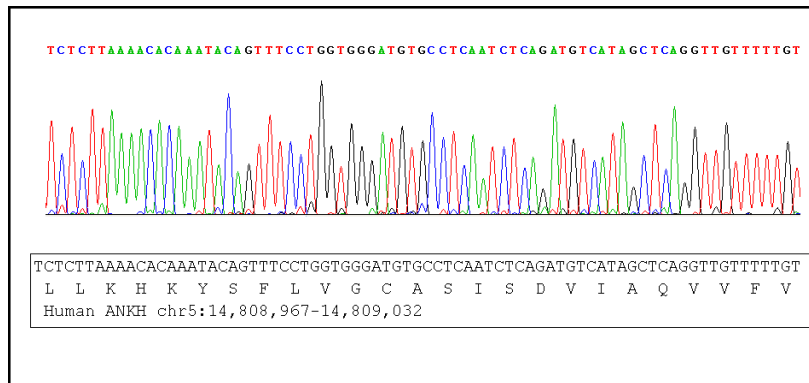
A

L HEK IMAGE K C mCCD mK

500 bp →

B

TCTCTTAAAACACAAATACAGTTTCCTGGTGGGATGTGCCTCAATCTCAGATGTCATAGCTCAGGTTGTTTTTGT

TCTCTTAAAACACAAATACAGTTTCCTGGTGGGATGTGCCTCAATCTCAGATGTCATAGCTCAGGTTGTTTTTGT
L L K H K Y S F L V G C A S I S D V I A Q V V F V
Human ANKH chr5:14,808,967-14,809,032

Figure 3.4 Expression of an ANK amplification product in human kidney cell line and tissue

(**A**) RT-PCR of an internal (490bp) ANK product. mRNA expression is observed in human embryonic kidney (HEK293) cells (**HEK**) and human renal cortex (**K**). IMAGE clone 3927236 cDNA (Geneservice) containing human full length ANK coding sequence (**IMAGE**) is the positive control. The water control (**C**) and ladder (**L**) in 100bp steps are also shown. Comparison with results for the mouse collecting duct cell line mpkCCD$_{cl4}$ (**mCCD**) and mouse kidney tissue (**mK**) (run on a separate gel) are shown. RT negative and RNA negative controls have been omitted.

(**B**) Example of sequencing chromatogram for product from HEK293 cells. Identity of PCR products was confirmed by direct sequencing. Identical sequencing was observed for human renal cortex and IMAGE clone.
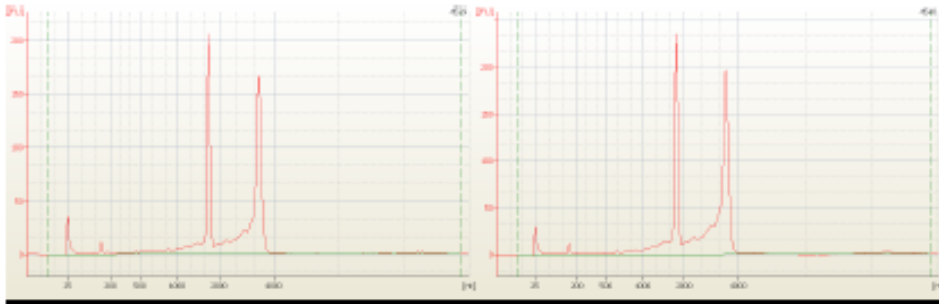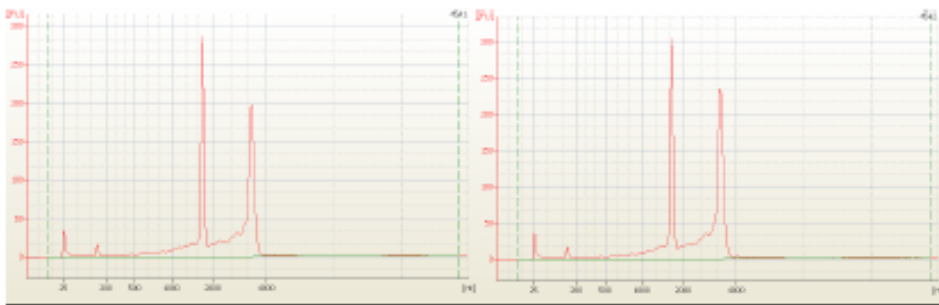
92

4537 C1 **4538** C2
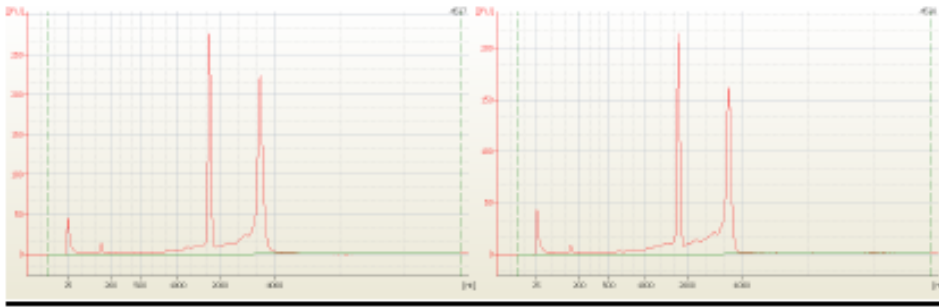
4539 C3 **4540** AVP1

4541 AVP2 **4542** AVP3

Sample ID	$\lambda(260/280)$	Concentration (ng/μl)	Yield (μg)
C1	2.03	399.6	3.2
C2	2.04	400.4	3.2
C3	2.04	431.8	3.5
AVP1	2.03	397.1	3.2
AVP2	2.04	505.0	4.0
AVP3	2.04	461.2	3.7

Figure 2.11 Validation of total RNA quality. Three control (C) and three AVP-treated (AVP) RNA samples were run on an Agilent bioanalyzer 2100. All results show two high and well-defined ribosomal peaks (28S, 18S). The small additional peaks below the ribosomal bands indicate a tiny amount of RNA degradation. The lack of peaks after the 28S ribosomal peak indicates a lack of genomic contamination. The table shows results of spectrophotometric analysis.
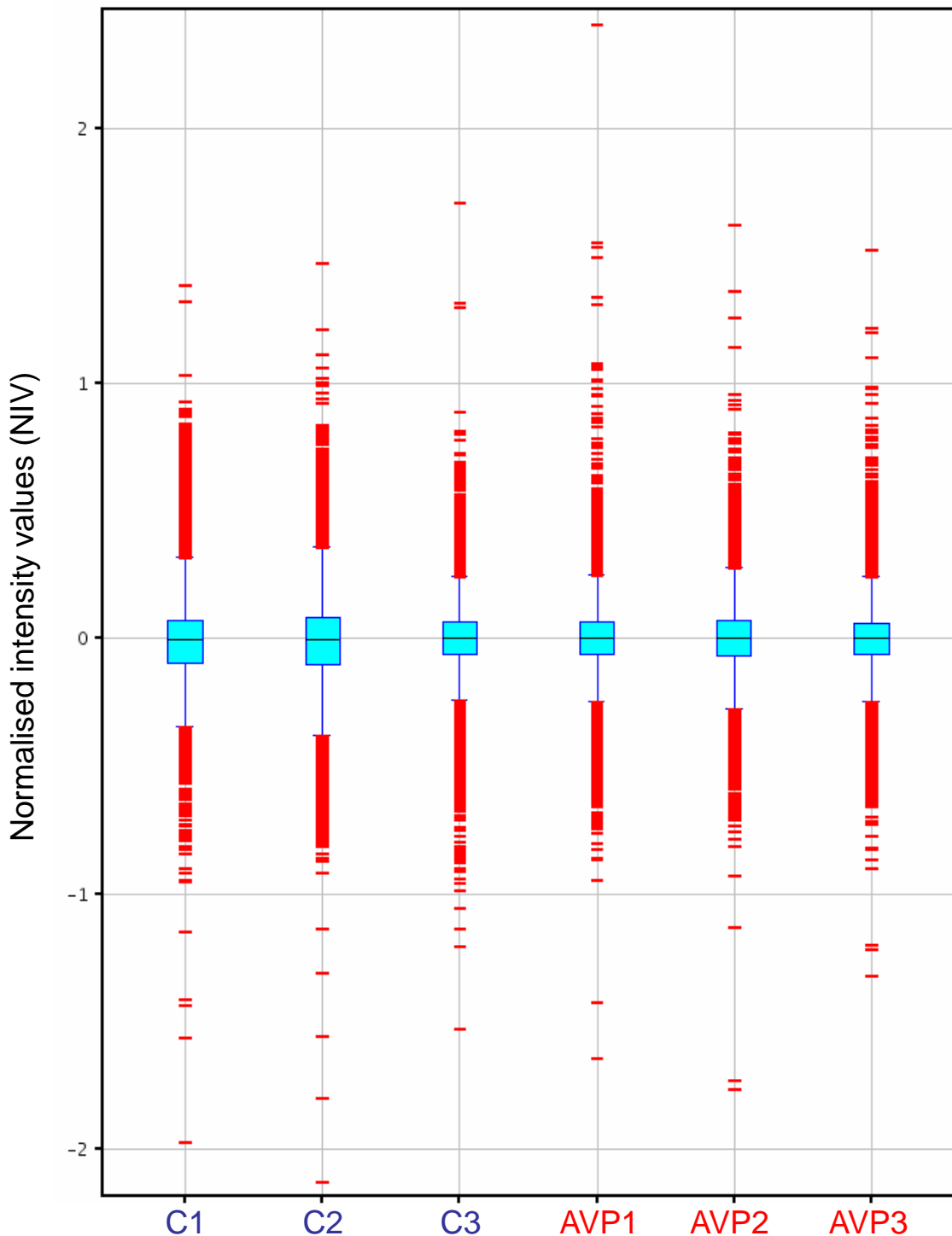
76

Figure 2.12 Normalisation of raw data. The log summarised values for each probe across all six samples is subtracted from each probe at each sample, hence reducing background 'noise' in the data. The resulting normalised intensity values (NIV) allow description of the position of each probe above the median (positive NIVs) or below it (negative NIVs). The data are seen here to distribute evenly across the experiment for all six samples, whether control (C) or treated (AVP).
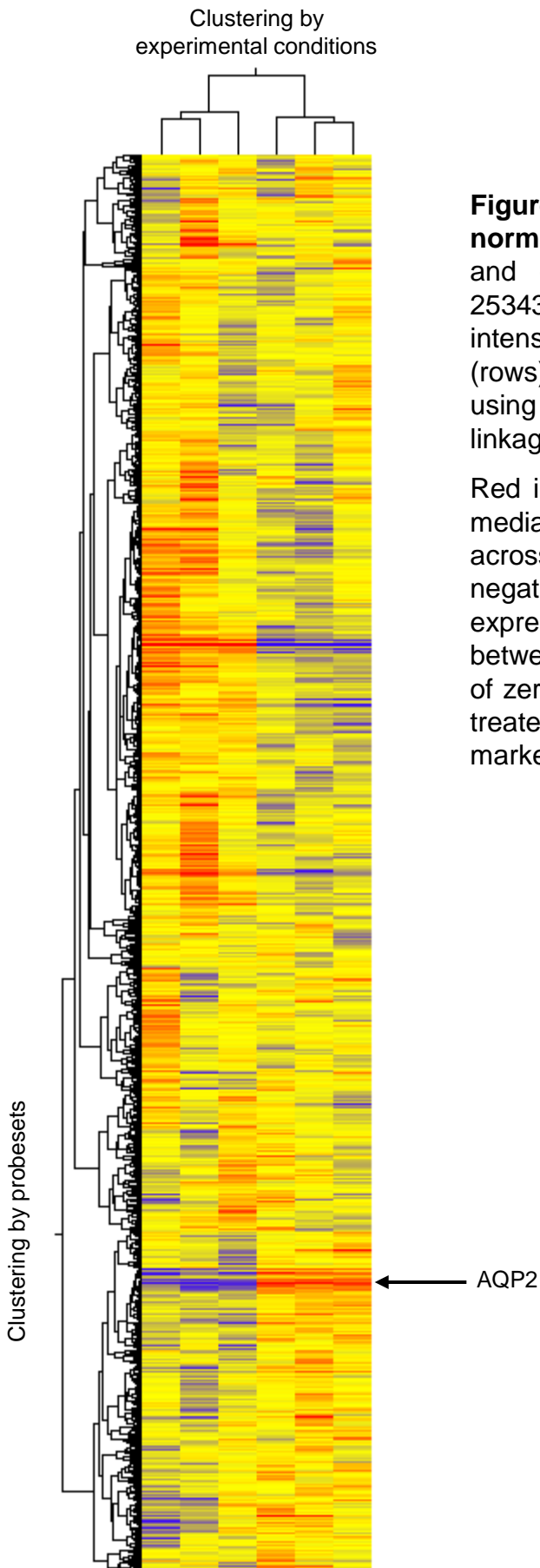
Figure 2.13 Hierarchical clustering of normalised microarray data. The full normalised and quality-controlled probeset list (containing 25343 probesets) was clustered by normalised intensity value (NIV) according to both probesets (rows) and control versus treatment (columns), using a Pearson centred correlation with average linkage.

Red indicates positive NIVs maximally above the median expression for that particular probeset across all six arrays, blue similarly indicates negative NIVs maximally below median expression. Shades of colour reflect NIVs between these extremes. Yellow denotes an NIV of zero, indicating no change between control and treated conditions. A probeset for aquaporin-2 is marked.
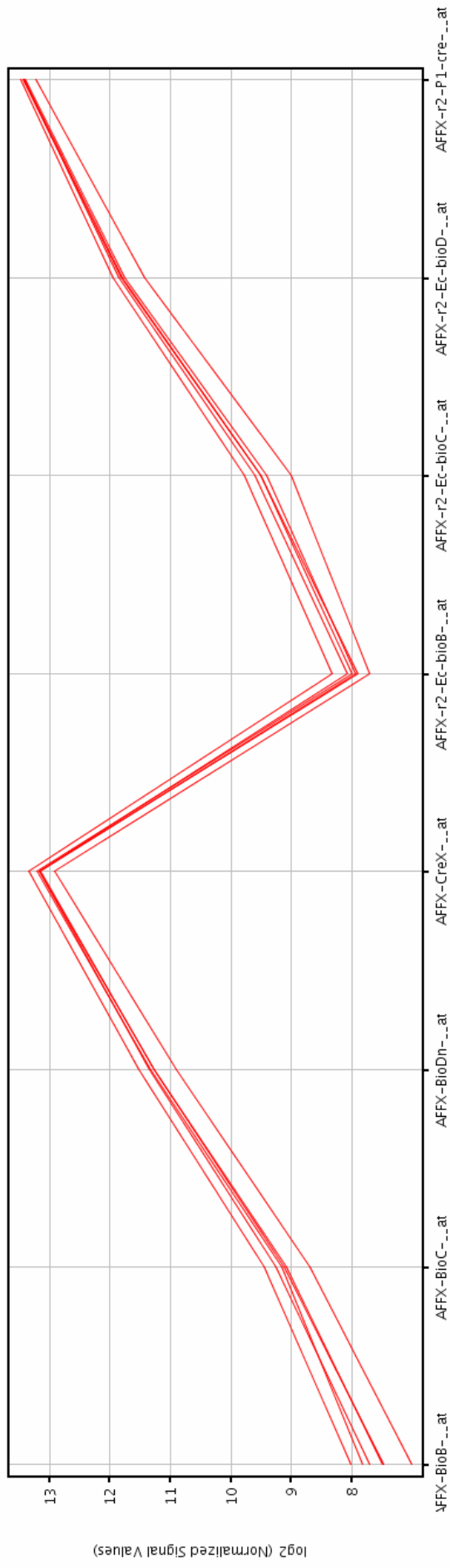
Figure 2.14 Hybridisation controls forming a 'spike' plot. Four additional RNA transcripts were added to the hybridisation mixture at known and increasing concentrations. The first transcript, bioB, is at the lower limit of detection of the arrays and has a 70% chance of being detected by the probesets (two for each transcript, causing the right-hand spike to look similar to the left-hand spike). The other controls (bioC, bioD, and cre) are well above the threshold for detection and should give rise to increasing normalised signal (expression) values (shown log transformed on the y-axis), as is the case here. Non-detection or values lower than expected would suggest a failure of the hybridisation reaction or subsequent washing and staining procedure.
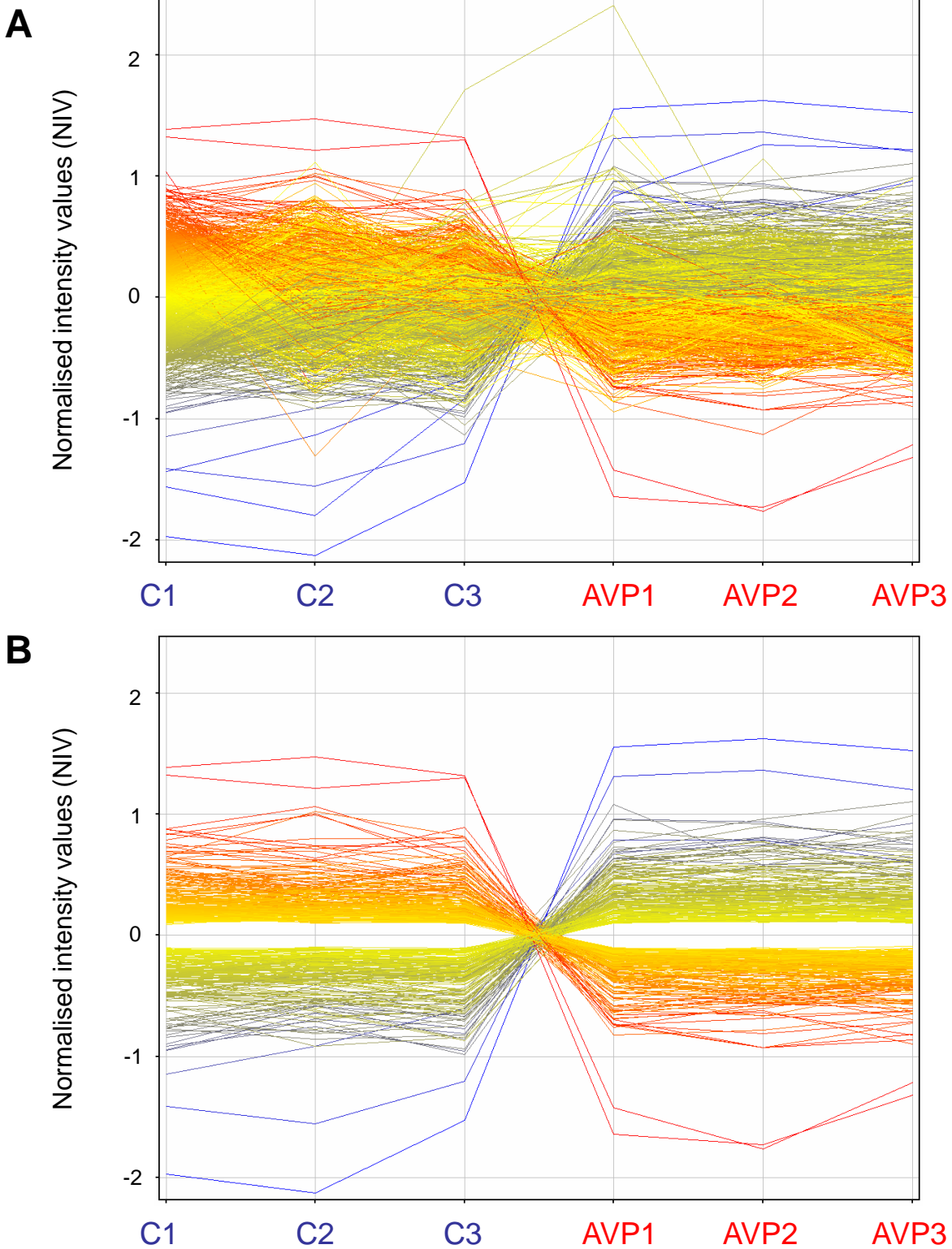
Figure 6.3 Change in expression of each probeset (gene) according to array. Each trend line connects a particular probeset at its normalised intensity value (NIV) in each of the six arrays (C1-3 control, AVP1-3 treated), with the crossover in NIV between conditions, if any, seen in the middle. An NIV of zero indicates the median value of the gene across all six arrays. (**A**) shows the raw data; (**B**) shows the final data (quality-controlled and t-tested). This comprises only the 364 probesets that changed by greater than 1.18-fold in either direction between the two conditions of the experiment, causing the central area of graph (**B**) to be blank. Detail of changes in specific genes are shown in Figure 6.6.
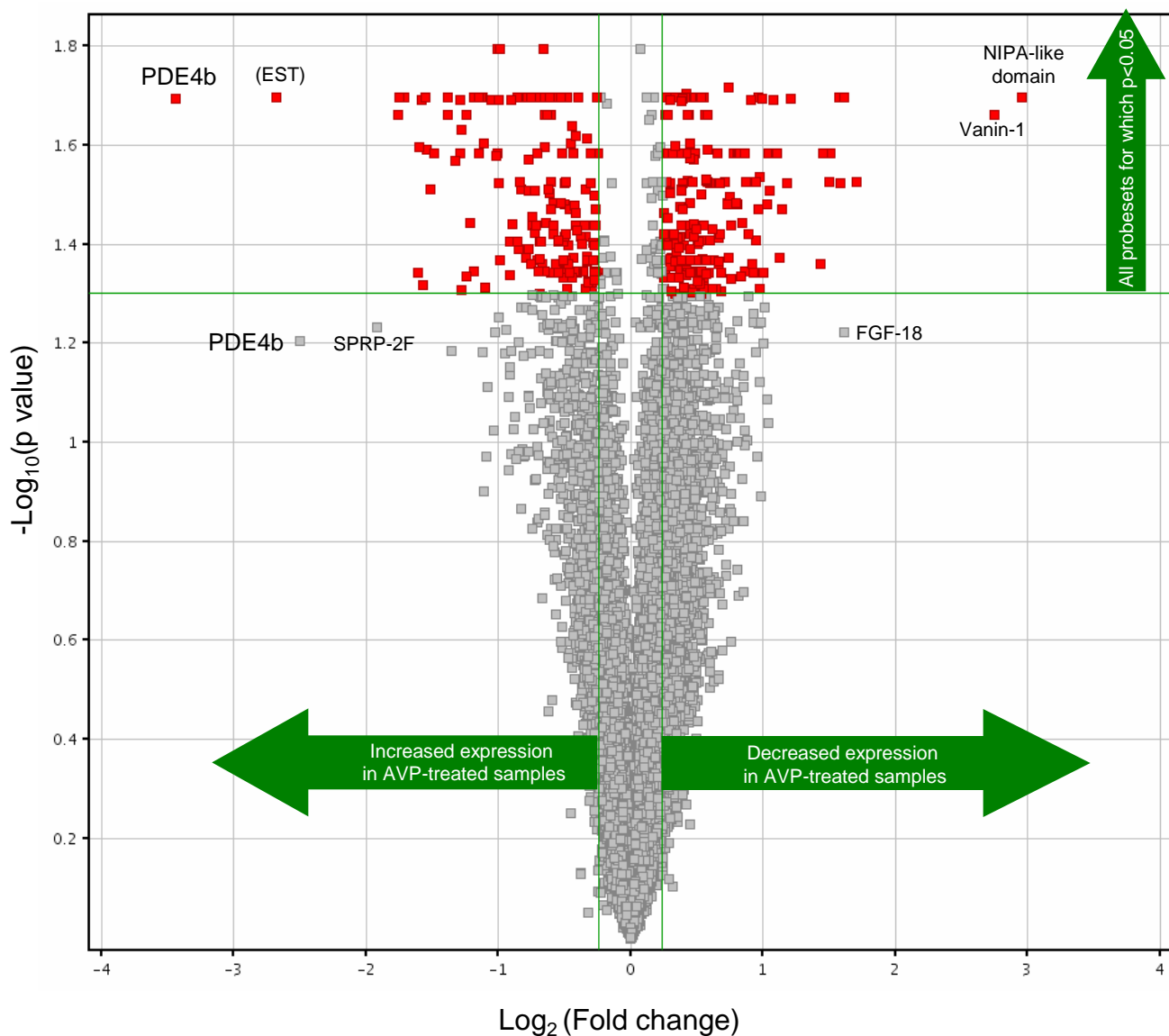
Figure 6.4 Change in probeset (gene) expression due to AVP treatment, for normalised and quality-controlled microarray data. All normalised quality-controlled data points are arranged by fold change of each gene in *control arrays with respect to AVP-treated arrays* (x-axis) and by p value (both expressed as log transformed values). Values above the horizontal green line represent those with $p < 0.05$, and values lateral to the vertical green lines represent fold changes of >1.18 in each direction. The combination of these two cut-offs is represented by the red values, which number 364 in total. Notable probesets with large fold changes are labelled. There were two probesets for PDE4b, only one of which was significant.

SPRP-2F: small proline-rich protein 2F; FGF-18: fibroblast-growth factor-18.

205

C1 C2 C3 AVP1 AVP2 AVP3
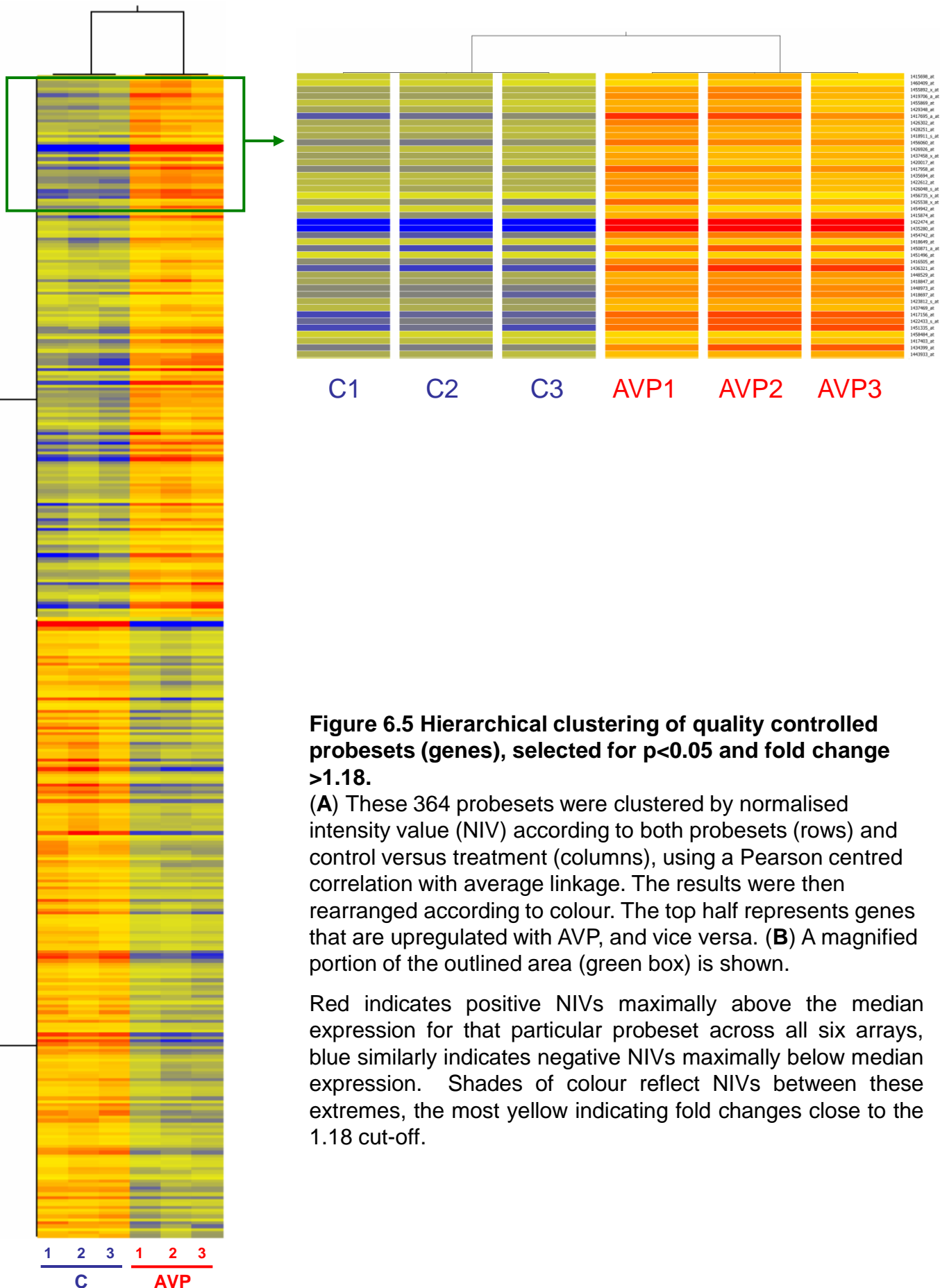
Figure 6.5 Hierarchical clustering of quality controlled probesets (genes), selected for p<0.05 and fold change >1.18.

(**A**) These 364 probesets were clustered by normalised intensity value (NIV) according to both probesets (rows) and control versus treatment (columns), using a Pearson centred correlation with average linkage. The results were then rearranged according to colour. The top half represents genes that are upregulated with AVP, and vice versa. (**B**) A magnified portion of the outlined area (green box) is shown.

Red indicates positive NIVs maximally above the median expression for that particular probeset across all six arrays, blue similarly indicates negative NIVs maximally below median expression. Shades of colour reflect NIVs between these extremes, the most yellow indicating fold changes close to the 1.18 cut-off.

1 2 3 1 2 3
C   AVP
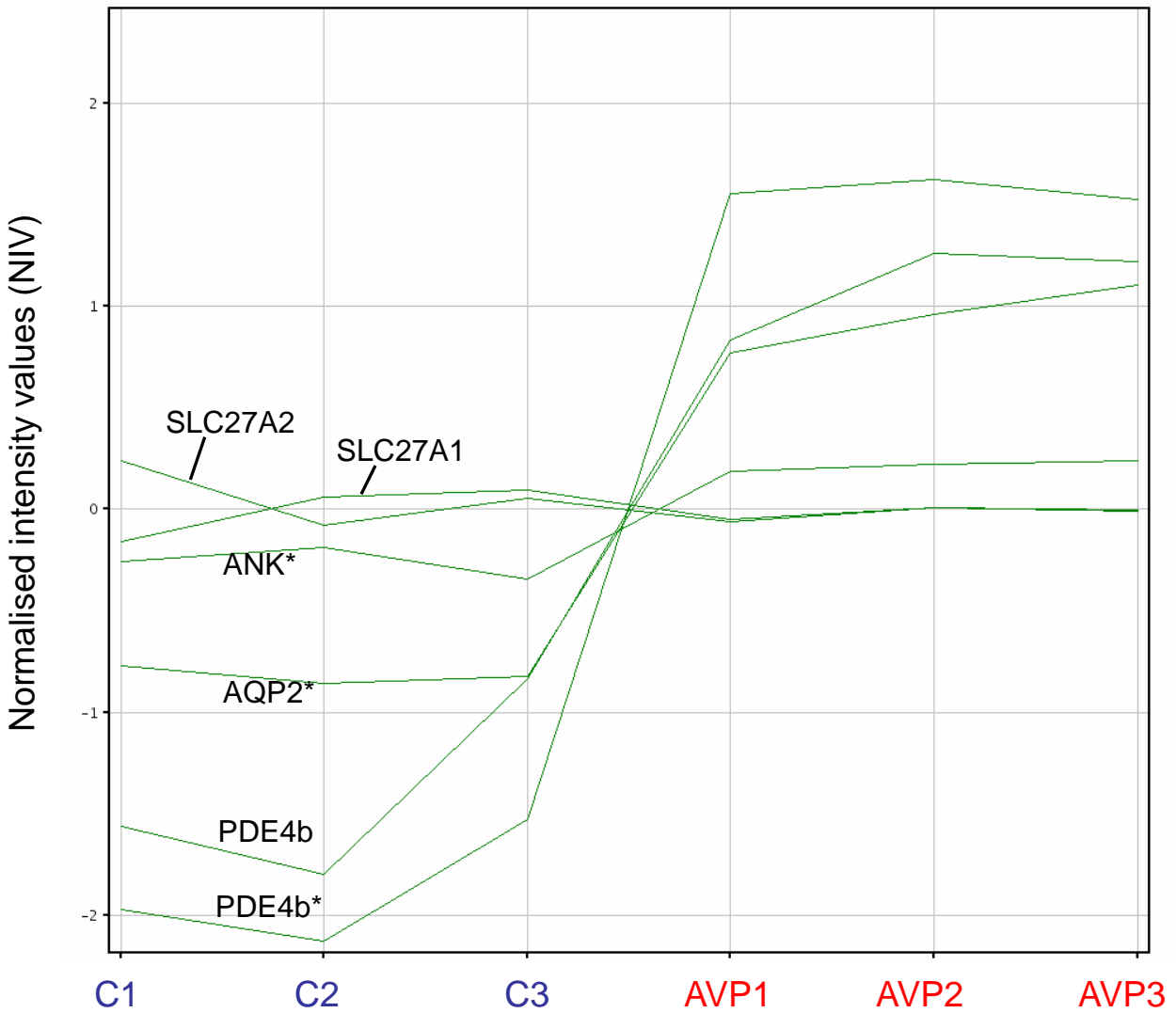
Figure 6.6 Change in expression of five genes with AVP treatment. Each trend line connects a particular probeset plotted at its normalised intensity value (NIV) in each of the six arrays (C1-3: control, AVP:1-3 treated). An NIV of zero indicates the median value of the probeset across all six arrays. Fold changes are given in Table 6.3. Results for both available PDE4b probesets are plotted, but only one of these (*) is significant. * denotes $p<0.05$.
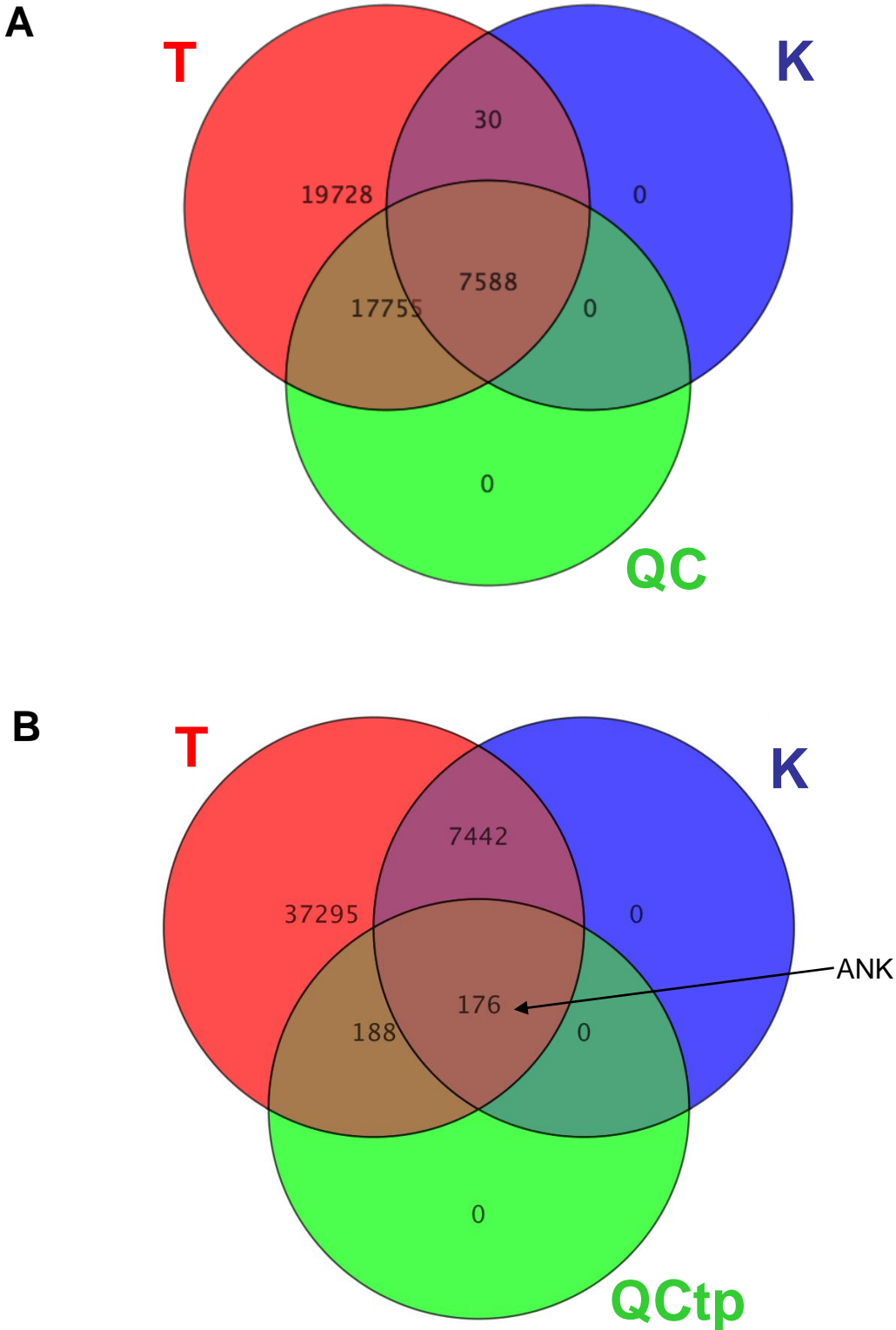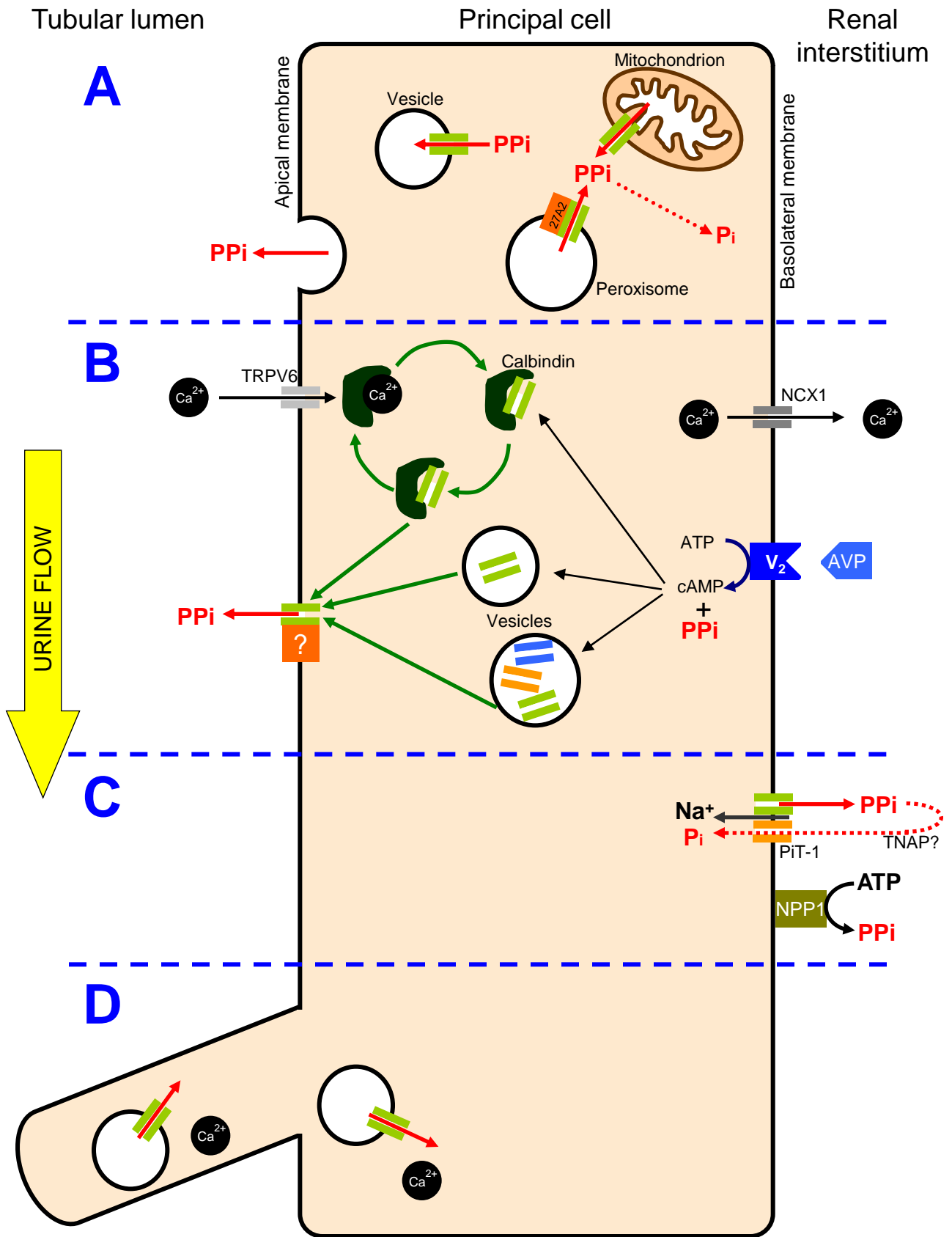
Figure 6.7 Summary of rationalisation of microarray data and overlap with Knepper dataset.

(A) The original 45101 probesets on the microarray (**T**) were reduced to 25343 probesets after normalisation and quality-control of the probeset list (**QC**). There was a high degree of overlap with the Knepper dataset (**K**), with only 30 genes of the 7618 in the Knepper dataset not being included in our QC list.

(B) The QC list was subjected to selection by t-test and fold change >1.18. This gives probesets that are differentially expressed with AVP treatment (**QCtp**). Only about half of these probesets were included in the Knepper list (**K**).

208

Tubular lumen | Principal cell | Renal interstitium

A

Vesicle — Mitochondrion — PPi

Apical membrane

PPi ←

27A2 — PPi ⋯→ Pi

Peroxisome

Basolateral membrane

B

Ca²⁺ — TRPV6 — Ca²⁺ — Calbindin — Ca²⁺ — NCX1 — Ca²⁺

URINE FLOW

ATP — V₂ — AVP

cAMP

PPi ← ? — Vesicles — + PPi

C

Na⁺ ← — PPi

Pi ← ⋯ TNAP?

PiT-1

ATP

NPP1 — PPi

D

Ca²⁺

Ca²⁺

Symbols — ⬛ ANKH — → PPi flux — → Movement of proteins or vesicles

236

AQP2

ANK

SLC27A2

Standards
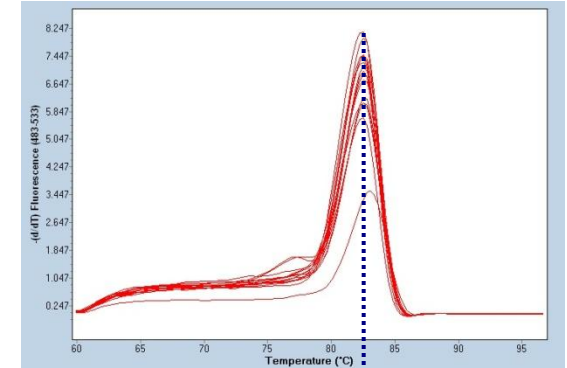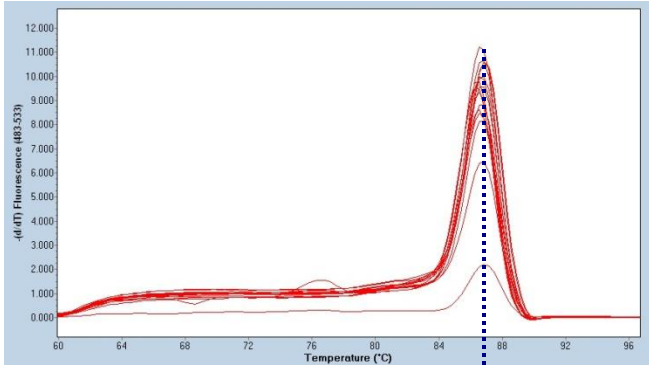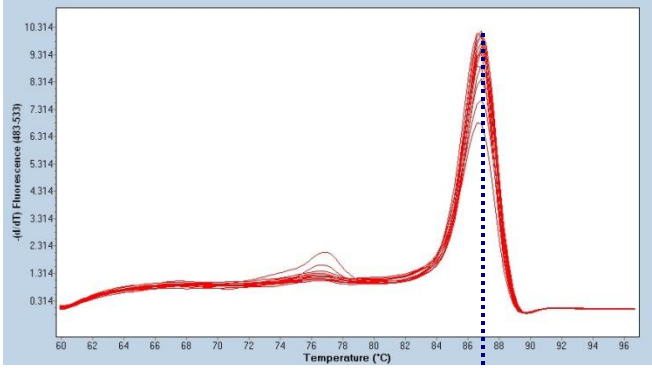
Experiment
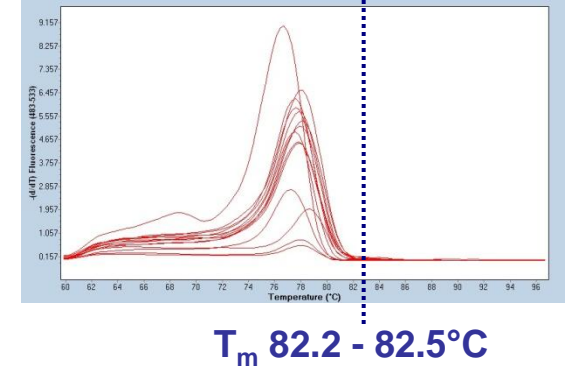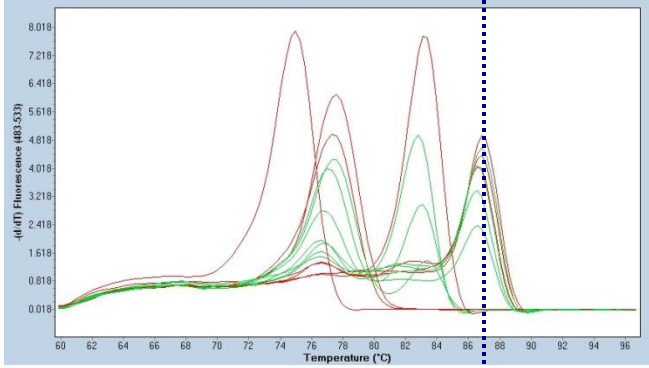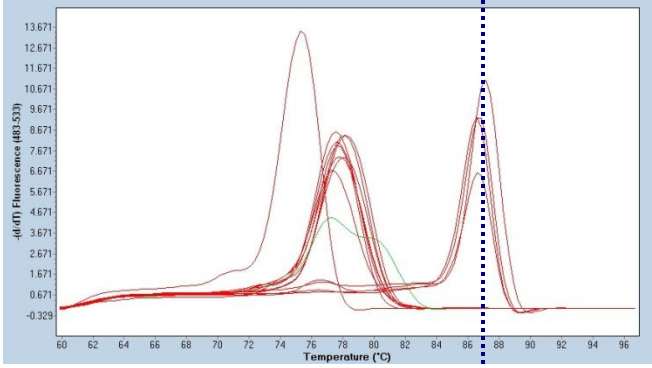
T$_m$ 86.5 - 86.8°C

(138 bp product)

T$_m$ 86.4 - 86.9°C

(89 bp product)

T$_m$ 82.2 - 82.5°C

(74 bp product)

L   A+   A-   C+   C-   X

75 bp →
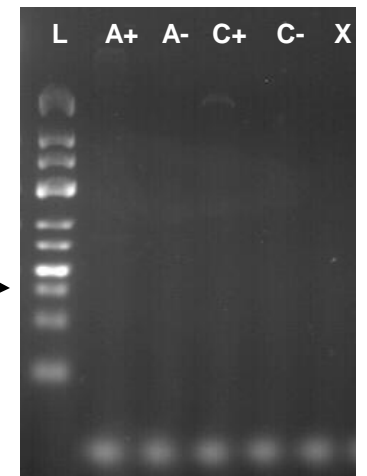
PDE4b

Efficiency 2.060

Slope -3.2
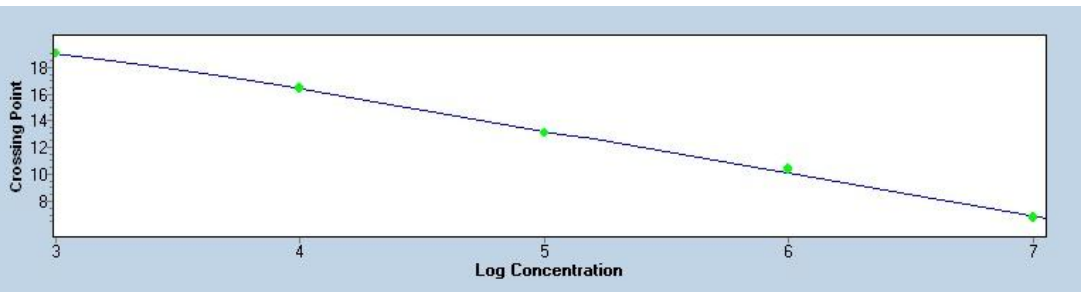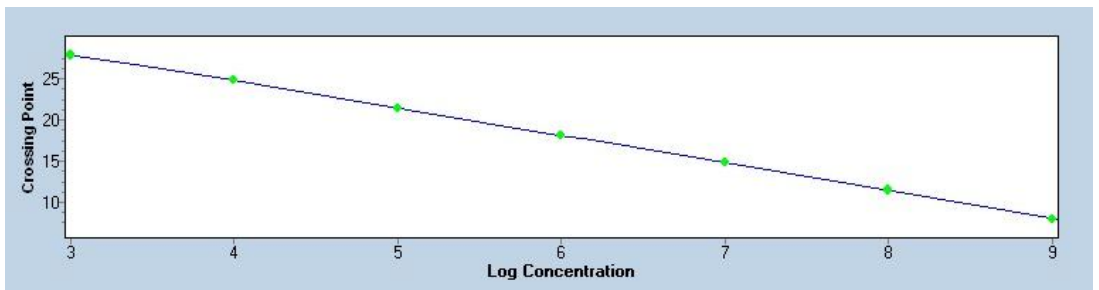

AQP2

Efficiency 1.918

Slope -3.5
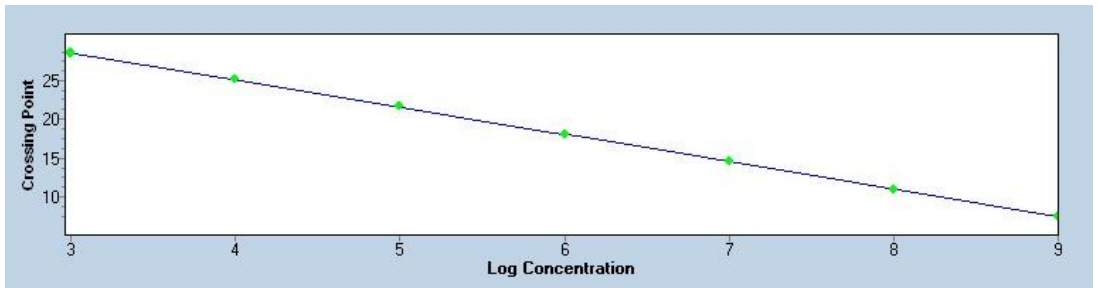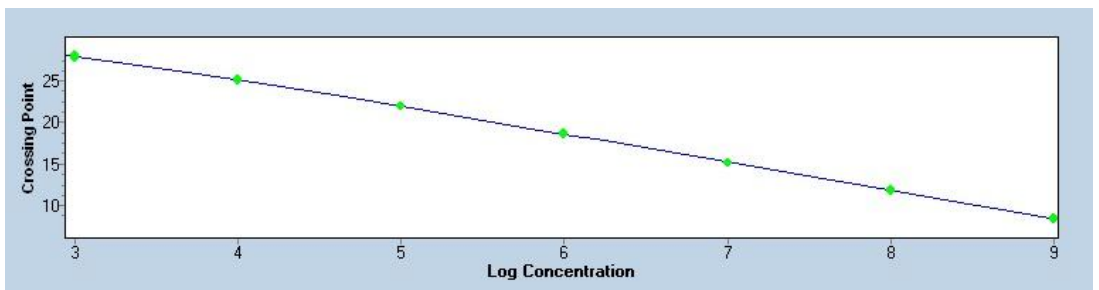

ANK

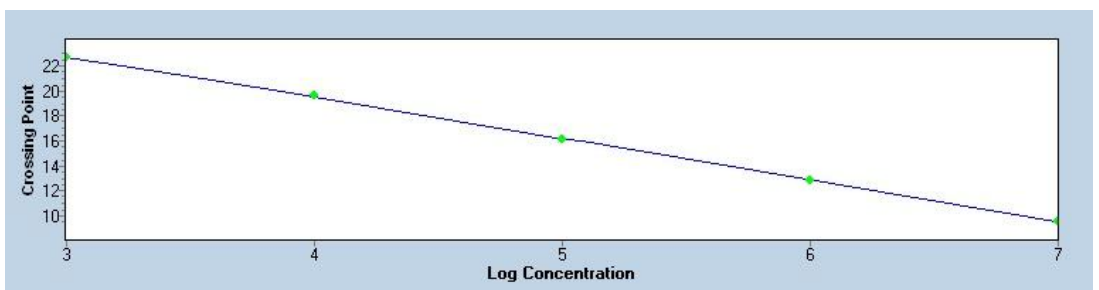Efficiency 1.992

Slope -3.3

GAPDH
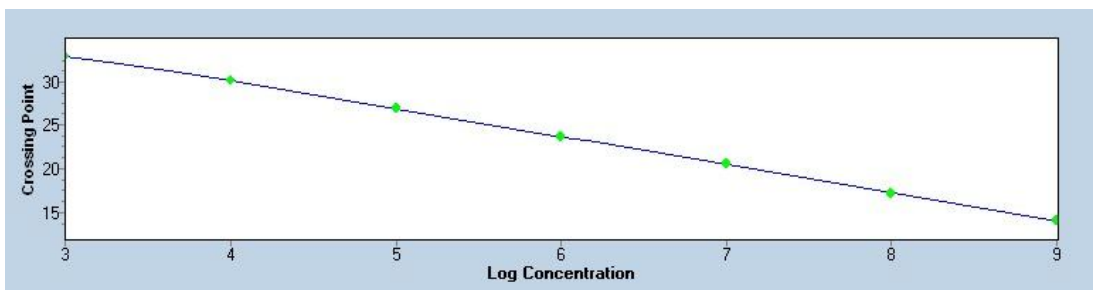
Efficiency 1.967

Slope -3.4

SLC27A1

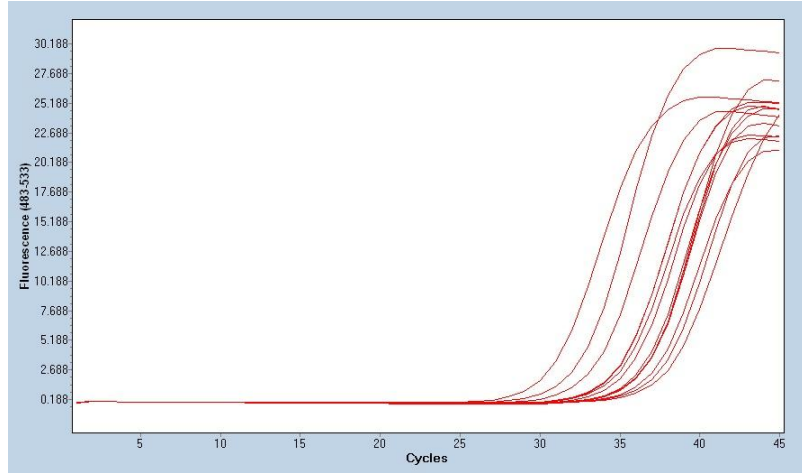Efficiency 1.985

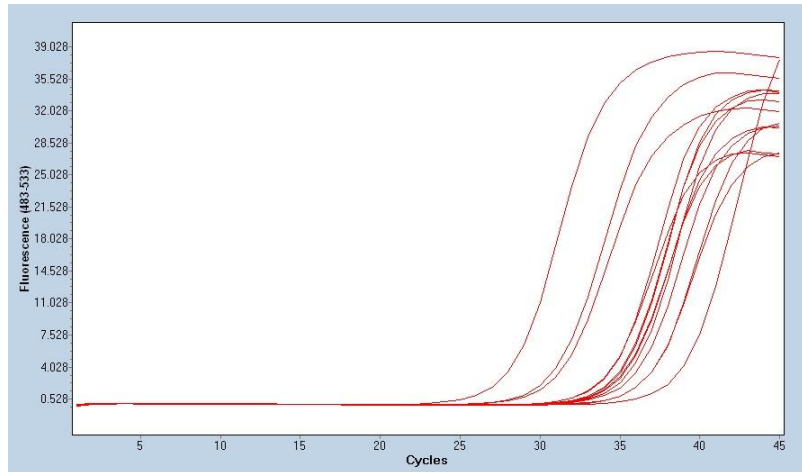Slope -3.3

SLC27A2

Efficiency 2.045

Slope -3.2

A

ANK standards

B

ANK

C

AQP2

D

GAPDH